Grand Challenge in MDC2

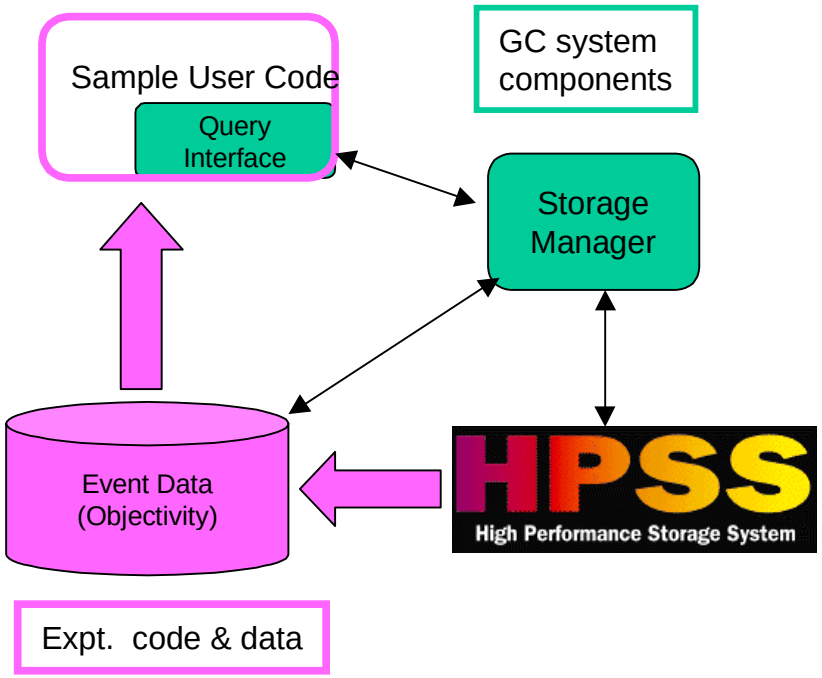
D. Olson, LBNL

31 Jan 1999

STAR Collaboration Meeting

http://www-rnc.lbl.gov/GC/

People

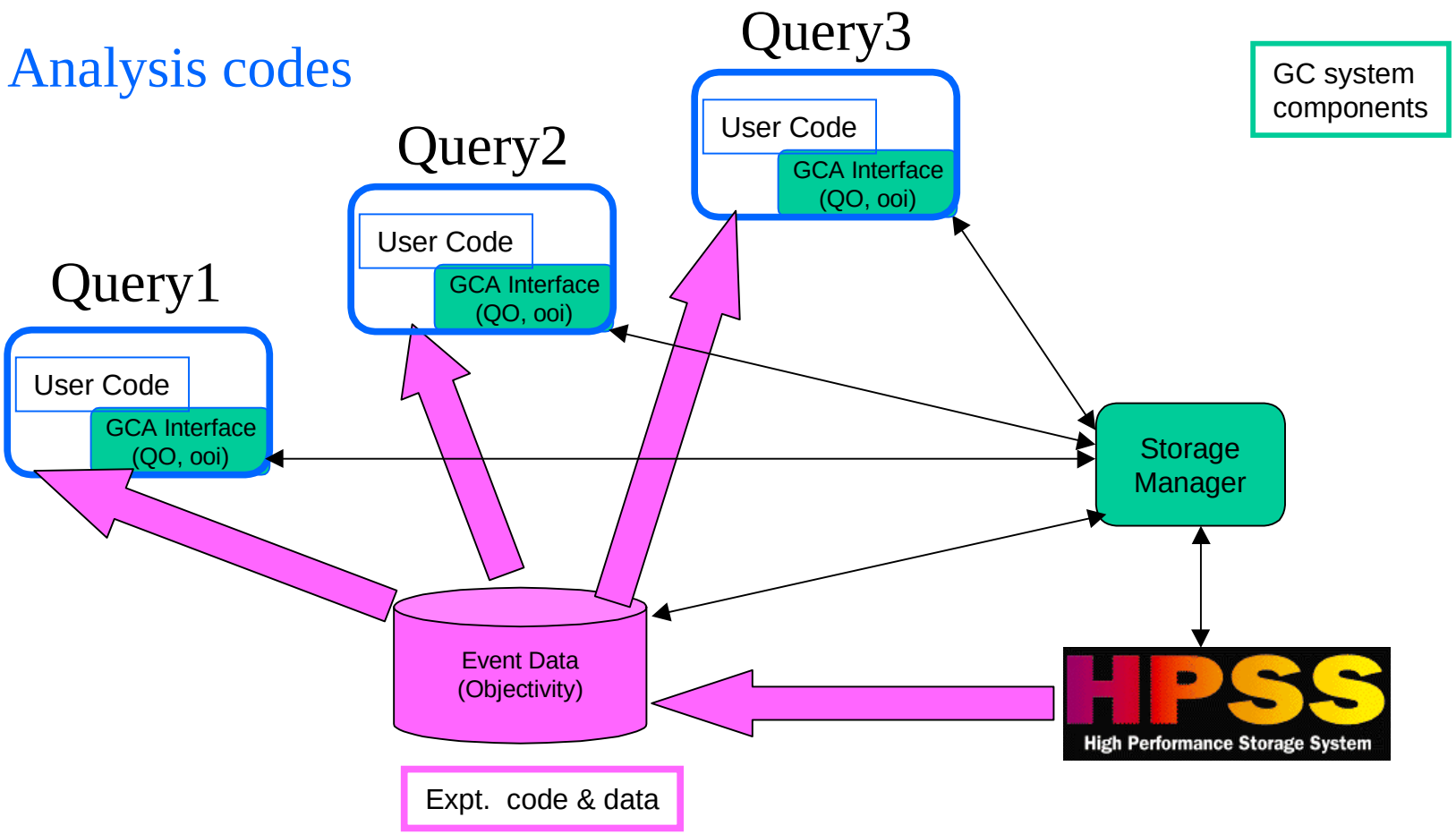


- key developers
  - Henrik Nordberg (NERSC) query estimator
  - Alex Sim (NERSC) query monitor
  - Luis Bernardo (NERSC) cache manager
  - Jeff Porter (BNL-STAR) query object
  - Dave Malon (ANL) order-optimized iterator & gcaResources API
  - Dave Zimmerman, (LBL-STAR) tagDB
  - Stephen Johnson (BNL/SB-PHENIX)
  - Jie Yang (UCLA,LBL) testing

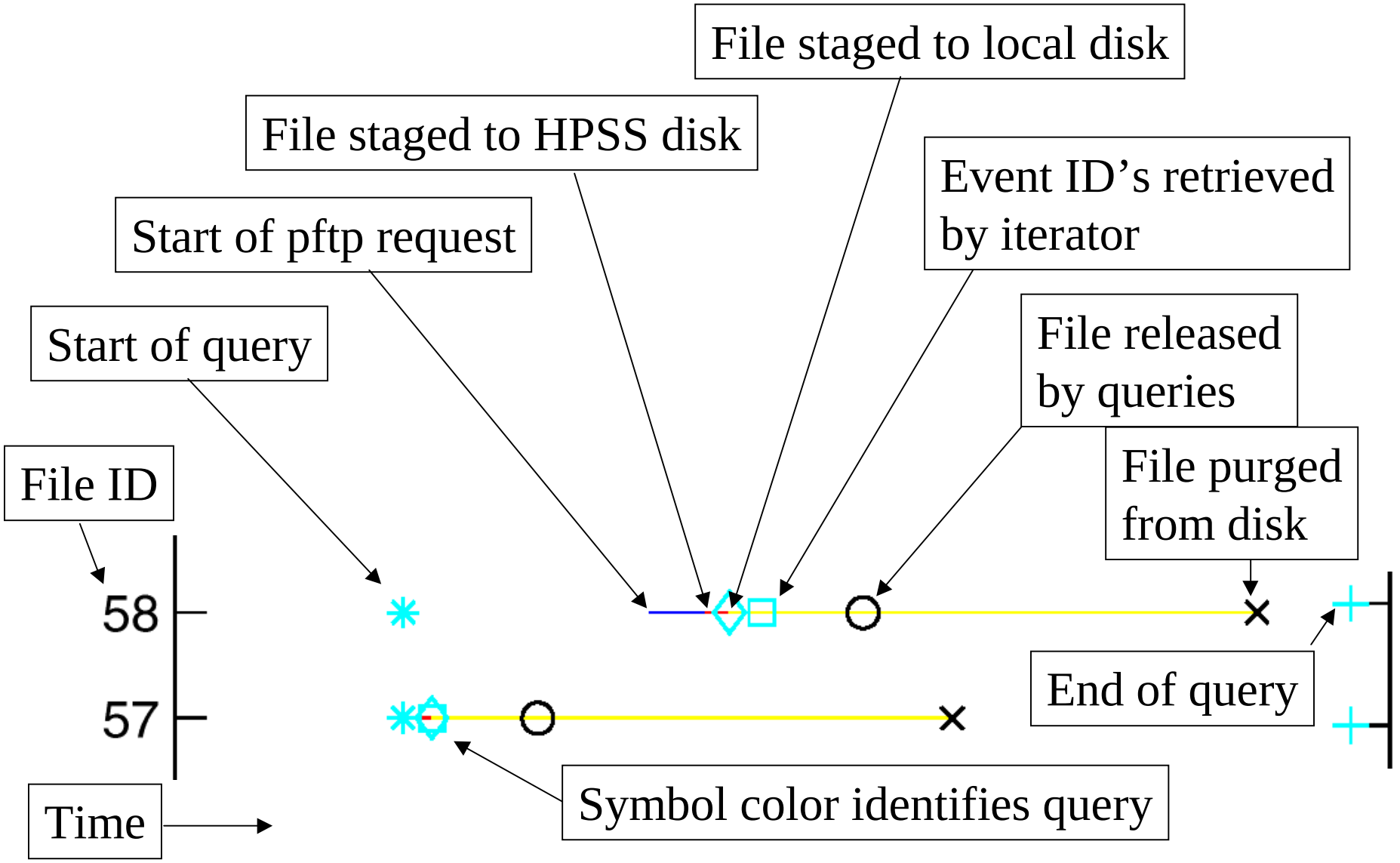
Outline

- Review of GC in MDC1
- MDC2 Features
- Status of Software
- Scenario of Usage
- Post-MDC2 (ROOT-GC integration?)
- Summary

Multiple simultaneous queries



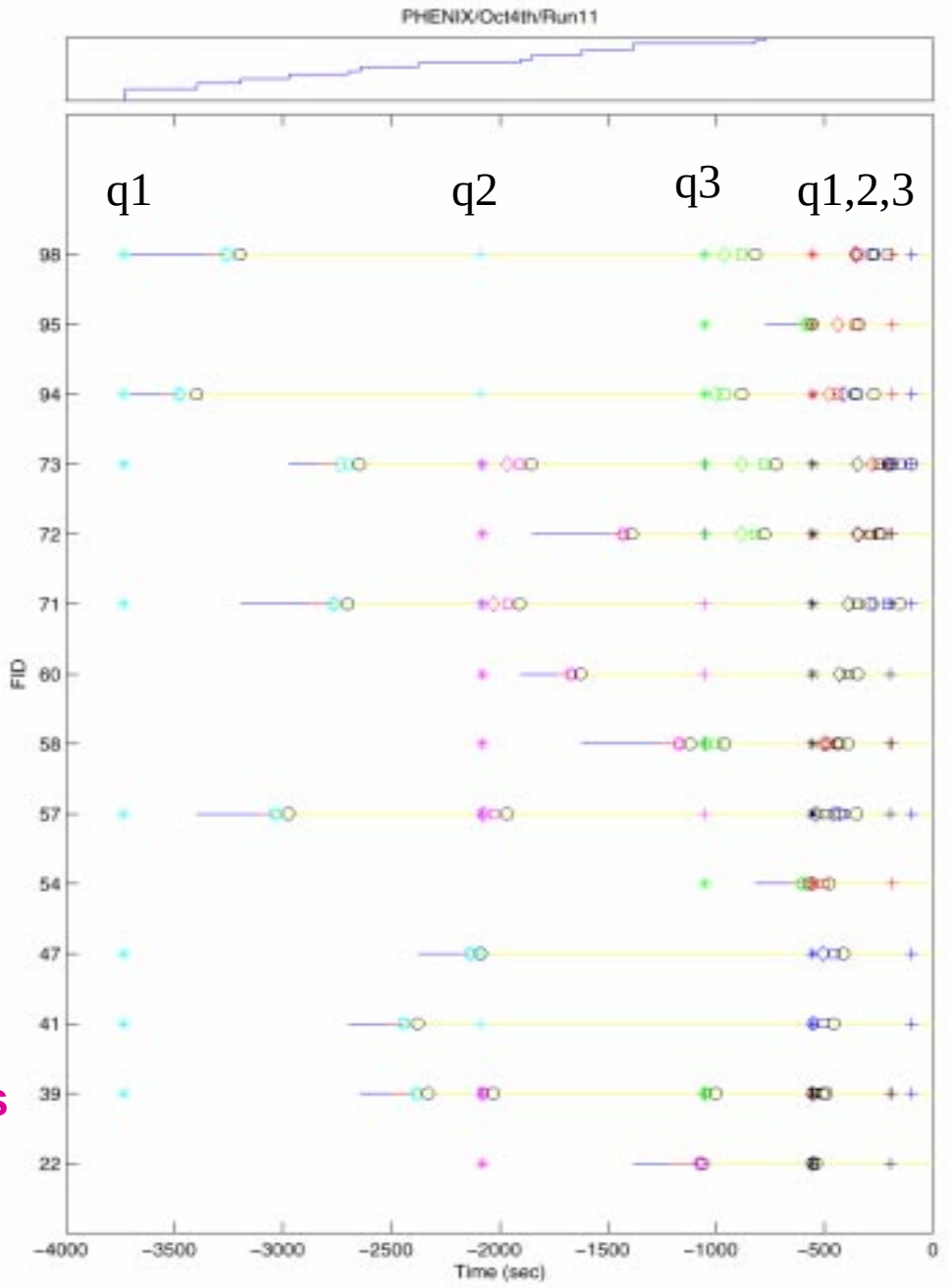
Legend


File staged to local disk
File staged to HPSS disk
Event ID's retrieved by iterator
Start of pftp request
Start of query
File released by queries
File ID
File purged from disk
58
57
End of query
Time
Symbol color identifies query

3 queries

3 queries with some shared files, time delay between each query, then the same 3 queries are repeated simultaneously.
The cache was large enough to hold all files so the second time all queries run at processing speed rather than I/O speed.

Detailed description of test results on GC web site under Results.

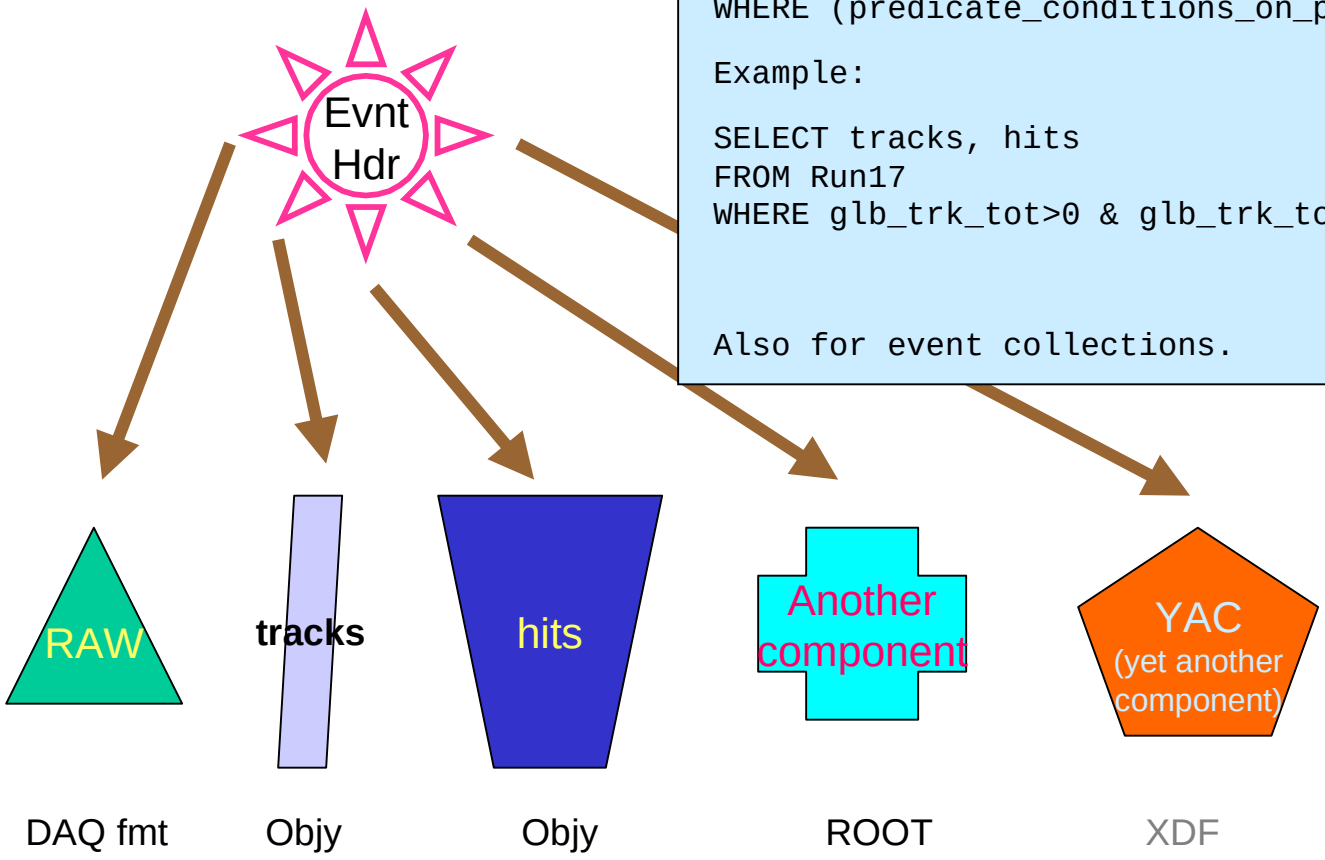


GC software features for MDC2

- Multiple event components
  - introduce “file bundles” - a set of events and their corresponding files
- Files in multiple directories (use Objy catalog)
- Time estimate for query
- parallel query execution (multiple iterators, same token)
- persistent state storage manager (can restart) (QE at least)
- query object can save event collection
- Linux analysis codes (Objectivity permitting) (PHENIX?)
- More logging
- See more details at http://www-rnc.lbl.gov/GC/email/archive/msg00663.html

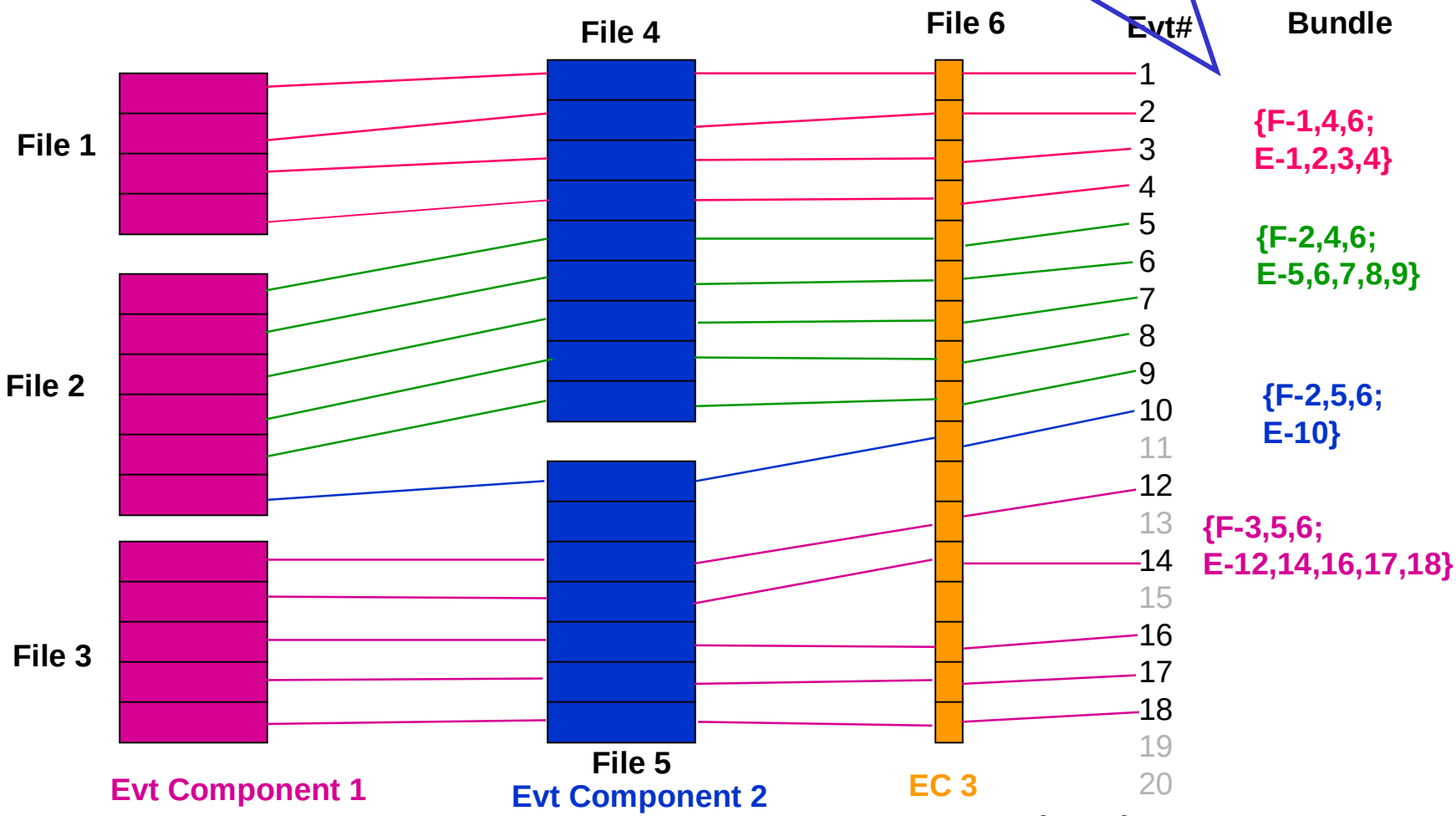
Multiple named components

```
SELECT (component_name, component_name, ...)
FROM dataset_name
WHERE (predicate_conditions_on_properties)

Example:

SELECT tracks, hits
FROM Run17
WHERE glb_trk_tot>0 & glb_trk_tot<10 & n_vert_total<3


Also for event collections.
```

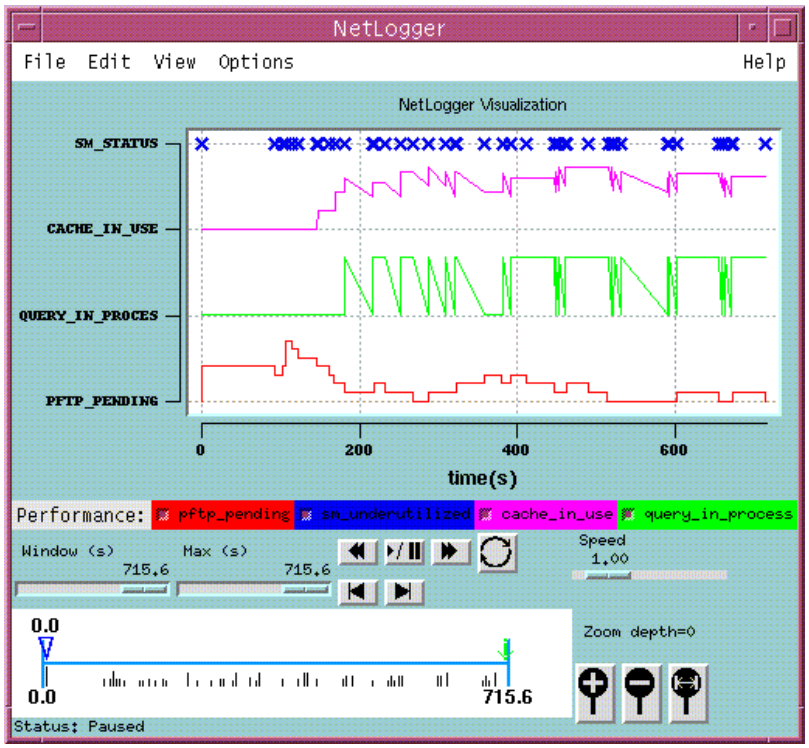


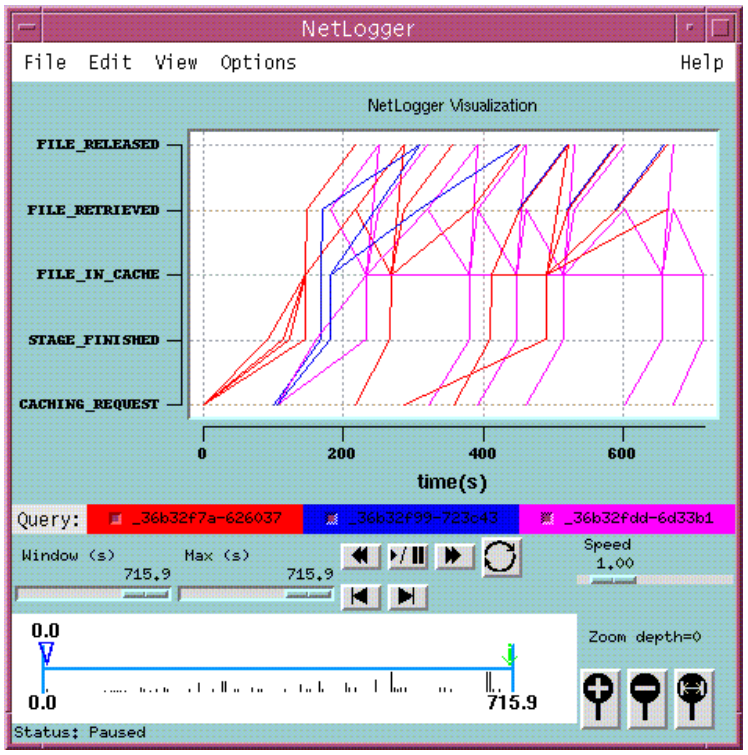
File Bundles

File bundles are delivered to the iterator in the analysis program.



Real-time Logging/Monitoring Example





Netlogger and Viewer from NERSC Data Intensive Computing Group (Tierney et al.)

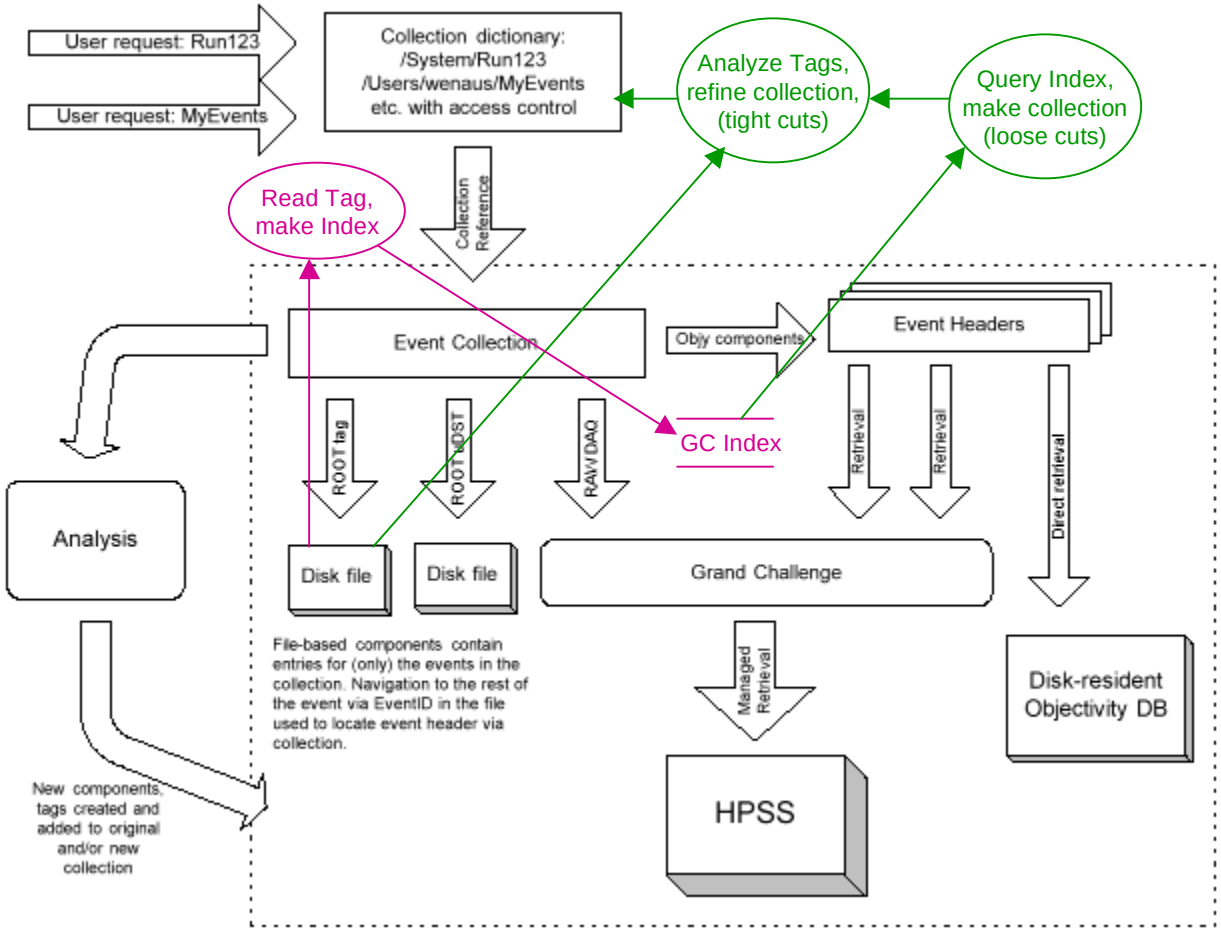
Status of s/w for MDC2

- Query Estimator - Ready
- Query Monitor - Done, Testing
- Cache Manager - Ready
- Logging - bug fixing
- Query Object - few mods required
- Order Optimized Iterator - few mods required
- ROOT-QO interface - proof of principle done, needs work
- ROOT TagDB interface to CG Index - Done

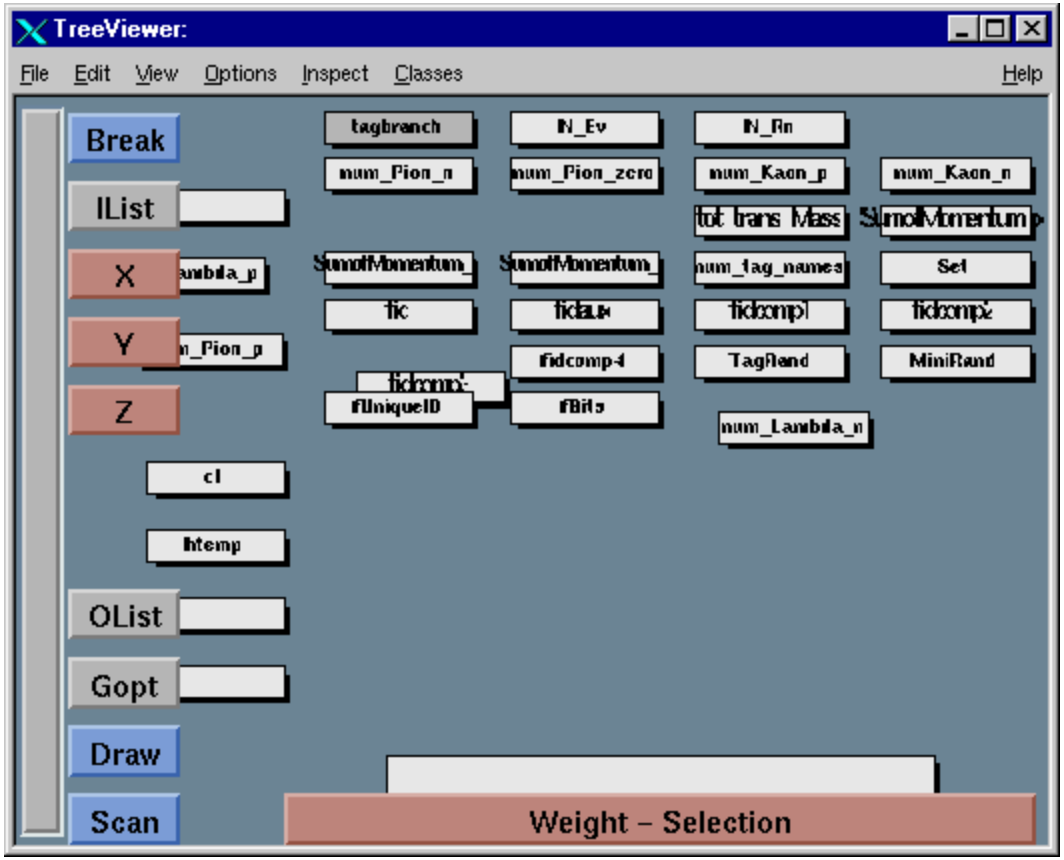
What needs to be done

- Interface with STAR event collections
- Interface with *StAnalysisChain*
  - OID --> build transient event
- Define content of STAR Tag Database in ROOT

Usage Scenario



TreeViewer of Tag Database in ROOT



Dave Zimmerman, Jan 31.
We will look at interfacing this to Query Estimator, GC Index.

Post-MDC2

- Work on scalability issues after MDC2
  - MDC2 does not stress scalability
- Tighter integration with ROOT
  - no Objy?
- Discussions w/ Rene Brun Jan. 25-27, 1999 at LBL

Topics w.r.t ROOT-GC

- PROOF
- TTree
  - TBranch
- TChain
- TEventList
- QO
- OOI
- Event ID’s
- TagDB
- Bit-mapped index

QO/OOI - TChain

- Iterating over the OID list retrieved from the storage manager for a file bundle is analgous to adding a file to a TChain and iterating over a TEventList for the selected events in that ROOT file.
- It would be possible to replace OID as event identifier with ROOT file + event sequence number in file in a non-Objectivity implementation.

What's To Do w.r.t. ROOT

- ROOT
  - mods to Chain
  - mods to EventList
- R.B. very interested in bit-mapped index
- GC
  - try OOI w/ Objy OID and get ROOT component w/ filename+evt seq.#
  - Consider non-Objy implementation (replace OID's)

Summary

- Basic MDC2 features ready.
- Need more work on the interface w/ STAR.
  - Any candidates? (LBL post-doc open)
- Need to clarify Objy/ROOT roles soon after MDC2.
- Plan on close integration w/ ROOT.